J. R. Statist. Soc. A (1990)
153, *Part* 3, *pp*. 321–347

Chance or Chaos?

By M. S. BARTLETT†

Exmouth, UK

[*Read before* The Royal Statistical Society *on Wednesday, March 14th, 1990, the President*, Professor P. G. Moore, *in the Chair*]

SUMMARY

In my survey of mathematical statistics in 1940 the concept of *chance* was assigned a central role. Some remarks are now made on the comparatively new concept of *chaos* (which can arise in non-linear systems even when 'deterministic'), first of all in relation to two papers in epidemiology by Schaffer (1985) and Olsen *et al*. (1988), and then more generally on the recognition and properties of chaos and its relation with chance.

Keywords: CHANCE; CHAOS; DETERMINISM; EPIDEMICS; NON-LINEARITY; PROBABILITY; SEASONAL VARIATION; STOCHASTICITY

1. INTRODUCTORY COMMENTS

About 50 years ago I presented a paper to the Royal Statistical Society (Bartlett, 1940) which surveyed the current situation at that time in mathematical statistics. A crucial role was assigned to probability theory, though this was to some extent challenged by A. L. Bowley and M. G. Kendall among the discussants (who in view of the world situation in 1940 made their contributions in writing); moreover, I emphasized that my interpretation of probability was to be closely associated with the concept of chance. To quote (p. 8):

> 'In practice we find that it is often possible to ascribe probability numbers to certain events, at least with as much justification as we can ascribe mass or temperature to physical systems. Events may be said to be governed in part by the operation of the "laws of chance".'

Since then, the post-war era has seen some obviously important developments that have affected the way in which statisticians model and analyse their observed data. One, with which I have been much involved, was the development of stochastic process theory to assist in the formulation of stochastic models in connection not only with the statistical analysis of time series but also with other applications of 'applied probability'. Another is the development of the modern computer, which has now become such an essential tool for the statistician. However, the computer has even more recently played a vital role in the development of what can be called (if it does not sound too contradictory) the *theory of chaos*, i.e. the theory which investigates

> '. . . the ability of even simple equations to generate motion so complex, so sensitive to measurement, that it appears random. Appropriately, it's called *chaos*.'

†*Address for correspondence:* Overcliff, 4 Trefusis Terrace, Exmouth, EX8 2AX, Devon, UK.

 0035–9238/90/153321 $2.00

(See Stewart (1989), p. 16.) May (1987), p. 31 (see also Stewart (1989), p. 269), has remarked, concerning its historical development:

> 'Given that simple equations, which arise naturally in many contexts, generate such surprising dynamics, it is interesting to ask why it took so long for chaos to move to centre stage the way it has over the past ten years or so. I think the answer is partly that widespread appreciation of the significance of chaos had to wait until it was found by people looking at systems simple enough for generalities to be perceived, in contexts with practical applications in mind, and in a time when computers made numerical studies easy.'

It is evident that statisticians concerned in any way with model building and the use of models for statistical applications have the responsibility of studying and trying to understand *chaos*, in various fields of application. I am not so presumptuous to think (in my 80th year) that I am the person to try to pronounce in any authoritative way on this unfamiliar theoretical discipline, but what I can do, with the help of some rather fragmentary comments, is to illustrate its relevance, raise some questions and at least indicate its interest and importance to statisticians, whatever their particular field.

Let me stress again that I am not attempting here any general exposition of this fascinating topic, which has relevance to all the sciences from physics to biology, including meteorology and economics. Those of you not already familiar with it should refer to the writings of some of those who have been most active in its development; see, for example, the 1987 Royal Society discussion on *Dynamical Chaos* (Berry *et al.*, 1987), which includes the paper by May already quoted, the book on *Non-linear Dynamics and Chaos* by Thompson and Stewart (1986), or the somewhat more popular books by Gleick (1987) and Stewart (1989).

2. CHAOS AND EPIDEMIOLOGY

Some of my more specific comments will relate to the field of epidemiology, as this subject is one with which I have been involved, although 30 years or more ago now (e.g. Bartlett (1949, 1956, 1957, 1960)). My own work then emphasized the relevance of stochastic models, e.g. in ecology and epidemiology, including in particular studying the properties of a stochastic version of what I shall still call for convenience the Hamer–Soper model (see Hamer (1906), Soper (1929) and Bailey (1957)) for recurrent measles epidemics, though Dietz and Schenzle (1985) have drawn attention to independent early work by En'ko (1889) and Martini (1921).

In contrast, let me quote from a comparatively recent paper by Schaffer (1985), where in his introduction he remarked:

> '. . . it is interesting to consider the possible consequences to ecology and epidemiology were it the case that nonlinear methods applied, and further, if these methods revealed evidence for chaos.
>
> 'The obvious consequence is that fluctuations previously believed random would turn out to have a deterministic basis. This would send the theorists back to their equations.
>
> 'At the same time, the dynamical reference point i.e. the baseline dynamics assumed for the deterministic part of the motion, would change. In the specific case of childhood epidemics, the emphasis on cyclic pattern of infection e.g. biennial outbreaks in measles, three year cycles in pertussis, etc. (e.g. Anderson *et al.*, 1984) would prove misplaced. In ecology, the equilibrium view of the world (e.g. Lewontin, 1969) central to two decades of theorizing would go forever out of the window. In particular, it would no longer make

sense to think of such systems in terms of a balance between intrinsic forces, forever searching out some mythical attracting point, and environmental vagaries perturbing the system away from it. Similar thinking pervades population genetics (e.g. Wright, 1968–1978), most notably in the theory of adaptive topographies.'

This quoted passage envisages the possibility of a dramatic switch in theoretical thinking. How far does Schaffer justify this view? His paper included a detailed investigation of the 'dynamics' of the so-called SEIR epidemiological model, where susceptibles (S), infected, but not yet infective (E), infective (I) and recovered (and immune) individuals (R) are governed by the deterministic model (in continuous time):

$$\left.\begin{aligned} \mathrm{d}S(t)/\mathrm{d}t &= m\{1-S(t)\}-b\ I(t)\ S(t), \\ \mathrm{d}E(t)/\mathrm{d}t &= b\ I(t)\ S(t)-(m+a)\ E(t), \\ \mathrm{d}I(t)/\mathrm{d}t &= a\ E(t)-(m+g)\ I(t). \end{aligned}\right\} \tag{1}$$

(Equations (1) are Schaffer's equations (8), with an obvious misprint correction to his first equation.) The total population $S+E+I+R$ is assumed constant, with births balancing deaths at rate m. Equations (1) are similar to the deterministic form of the stochastic model that I proposed (Bartlett, 1949), with the inclusion of a constant births or deaths term m and the extra class of latent infecteds E. The weakly damped deterministic oscillations which were a feature of the original model are not drastically affected by the slight extensions in model (1), as noted by Schaffer (1985), p. 237.

An important further extension of model (1), however, is to allow the infectivity coefficient b to have a periodic component to represent seasonal effects, as suggested by Soper (1929); see also, for example, Bartlett (1956) and Dietz (1976). The two conclusions reached by Schaffer (1985) were (p. 249):

'The first, maintenance of otherwise damped oscillations by noise appears incapable of reproducing essential features of the data. The second, cycles and chaos sustained by seasonal variation in contact rates gives qualitative and quantitative agreement between model and observation.'

My comments on these conclusions, which in the field of epidemiology appear to support Schaffer's general contentions, are as follows.

(a) It is not clear what sort of 'noise' Schaffer introduced. My own model was precisely formulated as a natural stochastic version of the deterministic model (in population studies deterministic models can be criticized for their treatment of *integers* like $I(t)$ and $E(t)$, which may at times be small, as *continuous* variables). Other sources of noise may exist as well, such as random variation of parameters (random environments) or spatial heterogeneity.

(b) The seasonal variation in infectivity adopted by Schaffer is rather larger in magnitude than that originally proposed by Soper and myself (in Schaffer's equation (10) for b, b_1 and b_0 should be interchanged).

(c) Nevertheless it would seem that the interaction between seasonality and the natural periodicity of the non-linear model is a significant factor which led to Schaffer's epidemiological conclusions. For small populations, where the intrinsic stochasticity becomes relatively more important, and, for diseases like measles to actual pockets of extinction of the infection, I have emphasized (e.g. Bartlett (1956, 1957)) important combined effects of stochasticity and extinction on the observed epidemic pattern.

In view of my comment (c) I was particularly interested to read also the more recent paper by Olsen *et al.* (1988), for in this paper some of my above criticisms have been met. A Monte Carlo 'simulation' of model (1) for the diseases measles, chicken-pox, rubella and mumps was reported, in which the possible transitions were assigned proportionate probabilities for a timescale that is sufficiently small for the sum of the probabilities not to exceed unity. (This presumably approximates to my own stochastic model if the timescale is sufficiently fine.) To circumvent the possibility of extinction of the infectives an additional immigration term for infectives was also introduced.

Large (fictitious) populations of 1 million and 5 million were investigated, the first value being chosen to correspond to the size of Copenhagen (Denmark). Olsen *et al.* claimed broad agreement with observation and with Schaffer's earlier conclusion of the 'chaotic' nature of the SEIR model, except for chicken-pox, where the yearly period is dominant. For smaller populations, however, Olsen *et al.* noted also the relevance of spatial heterogeneity, and the degree of linkage between communities.

To recapitulate, these (and other) researchers have demonstrated the potential relevance of chaotic fluctuations in SEIR models for large populations, due apparently to the interaction of the non-linear dynamics with the periodic seasonal factor. This is consistent with the extensive phenomenological demonstrations of the existence of chaos in many scientific disciplines. We must be careful, however, to balance any tendency there might now be to discern chaos everywhere with the responsibility of assessing its role in conjunction with other possibly relevant factors in any system under study. In the field of epidemiology, other relevant factors are those which I have previously emphasized of stochasticity, spatial heterogeneity and possible extinction of infection, especially for smaller populations. There is an obvious problem in the amount of relevant detail that should be inserted in a model, and age structure is another factor that has by now received attention (e.g. by Anderson and May (1983)). It would still seem important to determine all the main factors that are most essential for relating any successful model with its observed biological counterpart (in the case of chicken-pox, see also, for example, my remarks in Bartlett (1960), section 7.4).

3. STRUCTURES OF CHAOS IN CONTINUOUS AND DISCRETE TIME

The occurrence of chaotic behaviour in non-linear deterministic systems has raised the formidable problem of recognizing and categorizing it; for example, the conventional spectral analysis of time series data may be inadequate (see Schaffer (1985)).

Methods to date centre on the study of paths in 'phase space', i.e. in the multidimensional space of the variables of the system. For example, a simple pendulum has a two-dimensional phase space in its displacement $x(t)$ and its velocity $\dot{x}(t)$. After initial transients have damped out, this phase space path depicts what is known as the attractor, and ultimate chaotic behaviour corresponds to what is termed a *strange attractor*, as distinct from an equilibrium point or a limit cycle.

Some general theoretical results are available. Thus it is known (e.g. Thompson and Stewart (1986)) that, for continuous time systems, chaotic final behaviour requires at least three-dimensional phase space. An important example is the non-linear pendulum with a periodical forcing term (cf. the epidemiological model of Section 2). Notice that this does have a three- (not two-) dimensional phase space, because the periodic forcing term varies with the time t, which necessitates a third phase space variable t. The study of such a system can be reduced to a two-dimensional plot by choosing the values $x(t)$ and $\dot{x}(t)$ at the periods T, $2T$, . . . of the forcing term—this is an example of what is termed a Poincaré section, after the great French mathematician (Poincaré, 1899) whose topological studies of such orbits at the end of the last century have only been appreciated fully in recent years. Statisticians are more likely to have heard of his treatise on probability (Poincaré, 1896).

The difficulty of studying the behaviour of actual systems in multidimensional phase space is sometimes complicated further by the unobservability of all the relevant variables. A theorem of Takens (1981) (see Stewart (1989), p. 185) states that this problem can be bypassed by studying the multivariate time series $x(t)$, $x(t+h)$, $x(t+2h)$, . . ., where $x(t)$ is any one of the original variables of the system (and h is to some extent arbitrary). The path traced out by this concocted multivariate time series provides a topological approximation to the shape of the attractor.

This seems to provide a powerful technique for studying and detecting chaotic behaviour, but I imagine that much more investigation is needed to check the efficacy of the technique, especially when other sources of random behaviour may be present.

Note also that discrete-time systems can exhibit chaos even in one dimension. Thus the iterative equation

$$x_{n+1}=F(x_n) \tag{2}$$

that has sometimes been used as a model in ecology (e.g. Moran (1950)) can for quite simple functions $F(\cdot)$ exhibit complicated and sometimes chaotic behaviour, as stressed by May (1976, 1987). A well-known and remarkable example (e.g. May (1976) and Stewart (1989)) is the discrete-time 'logistic' iterative model

$$x_{n+1}=kx_n(1-x_n), \tag{3}$$

where x_n can represent population size for generation n (x being scaled to have maximum value unity) and k is a parameter between zero and 4. For $1<k<3$, x_n tends to a stable non-zero value $(k-1)/k$; for $k>3$, x_n becomes periodic for a further interval of k-values. Around $k=3.58$, however, chaos ensues (except for occasional 'windows' of particular k-values generating more regular periodic behaviour).

4. A CURIOUS EXAMPLE

These non-linear discrete-time systems might seem to promise easier classification and detection than continuous time systems, but there are curious features that need to be noted. Consider, for example, the simple autoregressive time series (linear stochastic model)

$$X_{n+1}=\rho X_n+Y_{n+1}, \tag{4}$$

where Y_{n+1} is independent of X_n and $\{Y_n\}$ is an independent series of values from some distribution. The correlogram is $\rho_n=\rho^n$, and the equivalent spectral density

$$f(\omega)=(1-\rho^2)/\pi(1+\rho^2-2\rho\cos\omega), \qquad 0<\omega<\pi.$$

If Y_n is normal (Gaussian), so is X_n. When X_n has become stationary, and, say, $E\{Y\}=0$, there are no further properties to discuss. The series reversed in time has identical statistical properties.

Consider next the stochastic model (4), with the same correlogram and spectrum, except that the random variable Y_n is not normal, but takes a few discrete values. The simplest case is two values, say $Y=0$ or $\frac{1}{2}$, each with probability $\frac{1}{2}$. To be a little more interesting, I shall take $Y=0$, $\frac{1}{3}$ or $\frac{2}{3}$, each with probability $\frac{1}{3}$. As n increases

$$X_n=Y_n+\rho Y_{n-1}+\rho^2 Y_{n-2}+\ldots. \tag{5}$$

The case $Y=0$ or $\frac{1}{2}$ will lead now to a uniform distribution for X, if $\rho=\frac{1}{2}$, as can be seen by our thinking of X, as expressed by series (5), as a random value between zero and unity written as a series on the binary scale. Similarly the case $Y=0$, $\frac{1}{3}$ or $\frac{2}{3}$ will lead to a uniform distribution for X, if $\rho=\frac{1}{3}$, as can be seen by expressing X on the 'triadic' scale. In the first case $E\{Y\}=\frac{1}{4}$ instead of zero, and in the second case $E\{Y\}=\frac{1}{3}$.

Thus, when X has become stationary, choose the initial value X at random between zero and unity; in the second case, say ξ when expressed on the triadic scale. For example, from the first vertical column of Table 33 (p. II) of Fisher and Yates's (1938) *Statistical Tables*, with (0, 1, 2) taken as 0, (3, 4, 5) taken as 1 and (6, 7, 8) taken as 2 (with 9 ignored), the first 10 numbers gave X_0 to sufficient accuracy as

$$\xi=0.1212022110\ldots$$

or 0.621 to three decimal places in ordinary decimal notation. The next number in the column gave 2, and this was taken to provide $Y=\frac{2}{3}$, so that on the triadic scale

$$X_1=0.1\times\xi+0.2=0.21212022110\ldots,$$

and so on. The 50 values $X_1,\ldots,X_{50}$ obtained in this way are given (to three decimal places) in Table 1 and are also shown in Fig. 1.

You may wonder what relation all this has to deterministic, let alone chaotic deterministic, behaviour. I think that it was Murray Rosenblatt who first mentioned this kind of example to me many years ago (see also Degn (1982)). If we consider the series *reversed in time*, it becomes transformed into a non-linear chaotic deterministic series.

The case illustrated in Table 1 and Fig. 1 may be reversed in time to read

$$x_n=3x_{n+1}-[3x_{n+1}], \tag{6}$$

where $[3x_{n+1}]$ denotes the 'integral part' of $3x_{n+1}$ on the triadic scale, or equivalently on the decimal scale. Clearly a plot of x_n, or even of the pair (x_{n+1}, x_n), would not automatically from its topology distinguish between the linear stochastic series (4) and the non-linear deterministic series (6), because, apart from the direction of time, they are identical. Plotting the pair (x_{n+1}, x_n) gives the regression relation between x_{n+1} and x_n; for Y Gaussian this would be contained in a bivariate Gaussian distribution, but for $Y=0$, $\frac{1}{3}$ or $\frac{2}{3}$ consists of three parallel straight lines with slope $\frac{1}{3}$ (two lines for the $Y=0$ or $\frac{1}{2}$ case), with of course the same average regression line as in the Gaussian case.

TABLE 1
Realization of linear stochastic model (4) (with $\rho = \frac{1}{3}$)

n	$3Y$	X	n	$3Y$	X	n	$3Y$	X	n	$3Y$	X	n	$3Y$	X
1	2	0.874	11	1	0.391	21	1	0.500	31	0	0.061	41	0	0.278
2	0	0.291	12	2	0.797	22	0	0.167	32	1	0.354	42	1	0.426
3	0	0.097	13	2	0.932	23	1	0.389	33	0	0.118	43	1	0.475
4	0	0.032	14	1	0.644	24	2	0.796	34	2	0.706	44	2	0.825
5	2	0.677	15	0	0.215	25	0	0.265	35	1	0.569	45	1	0.608
6	2	0.892	16	0	0.072	26	2	0.755	36	1	0.523	46	1	0.536
7	1	0.631	17	1	0.357	27	2	0.918	37	1	0.508	47	0	0.179
8	1	0.544	18	1	0.452	28	1	0.640	38	1	0.503	48	2	0.726
9	1	0.525	19	1	0.484	29	1	0.546	39	1	0.501	49	2	0.909
10	0	0.172	20	1	0.495	30	0	0.182	40	2	0.833	50	0	0.303

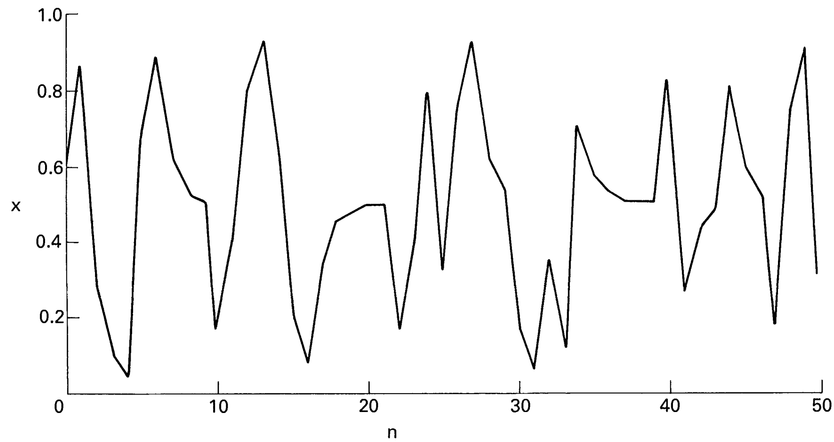

Fig. 1. Realization of stochastic model (4), with $\rho = \frac{1}{3}$ and $3Y = 0$, 1 or 2 (deterministic series (6) when reversed in time)

Three incidental remarks are as follows:

(a) if the structure of series (6) is disliked as too artificial, the 'integral part' term may be filtered off automatically by wrapping $3x_{n+1}$ round a circle (see Stewart (1989), pp. 112–113);
(b) from Fig. 1 it will be seen that (in contrast with the Gaussian case) the time direction is discernible from the shape of the fluctuations;
(c) the simple type of equation (6) will only be available for such special cases as $Y=0$ or $\frac{1}{2}$ with $\rho = \frac{1}{2}$, or $Y=0$, $\frac{1}{3}$ or $\frac{2}{3}$ with $\rho = \frac{1}{3}$, etc., these cases leading moreover to the uniform distribution for X, but some more arbitrary values of ρ could perhaps still be handled, though with more difficulty and with the aid of a computer.

A more relevant point is that the prediction problem, if confined to correlational properties and linear prediction, would be the same for equations (4) and (6), but a deterministic series, e.g. equation (6), should be predictable exactly, i.e. for a deterministic series (proceeding forward in time) x_{n+k} should be predictable precisely from x_n. However, not only must the equation for the system be precisely known,

TABLE 2
Prediction from the (reversed) deterministic series (6)†

n	*(a)*	*(b)*	n	*(a)*	*(b)*	n	*(a)*
49	0.909	0.909	44	0.825	0.887	39	0.484
48	0.726	0.727	43	0.475	(0.661)	38	0.451
47	0.179	0.181	42	0.425	(0.983)	37	0.352
46	0.536	0.543	41	0.276		36	(0.055)
45	0.608	0.629	40	0.828		35	

†(a) $x_{50}=0.302915$; (b) $x_{50}=0.303$.

but also the initial value x_n must be precisely known. The ultimate limitations to this last condition become obvious for series (6) when it is noted that as k increases the value of x_n as a triadic series needs to be known to k triadic places.

To illustrate this, the predicted values using the deterministic series (6) are given (to three decimal places) in Table 2, working back from the value x_{50}, when

(a) $x_{50}=0.302915$ to six decimal places and
(b) $x_{50}=0.303$ to three decimal places.

As anticipated, column (a) gives predicted values agreeing with x_n to within one in the third decimal place as far back as $n=42$, but breaking down rapidly by $n=38$–36 (about 12–14 steps), whereas column (b) only agrees to within one in the third decimal place back to $n=48$, and breaking down by $n=44$–42 (about 6–8 steps). This is a very elementary demonstration of the effect of the accuracy to which the 'initial' value of a univariate non-linear deterministic series must be known to ensure accuracy of the prediction so many steps ahead.

The accuracy to which x is recorded might be associated with an additional 'error term' ϵ; it then has some link with a general theoretical approach by Zeeman (1988) to a classification of the stability of dynamical systems, even chaotic ones, by introducing in the theory what he refers to as 'ϵ-diffusion'.

5. IMPLICATIONS FOR THE CONCEPTS OF PROBABILITY AND CHANCE

To return to the concepts of probability and chance which I mentioned in Section 1, we have seen that the 'strange' behaviour of chaotic systems has to be added to the possible forms of random behaviour with which some of us have been more familiar. The example in Section 4 has perhaps emphasized the need for care in discriminating between them. Extreme cases of non-linear chaotic behaviour which appear completely random are indeed already familiar in the pseudorandom sequences used in computers in place of random numbers. This links with the view expressed by Stewart (1989), pp. 299–300:

> 'Our most cherished examples of chance—dice, roulette, cointossing—seem closer to chaos than to the whims of outside events. So, in this revised sense, dice are good metaphor for chance after all. It's just that we've refined our concept of randomness. Indeed, the deterministic but possibly chaotic stripes of phase space may be the true source of probability.'

Such notions of probability are also illustrated by Galton's *quincunx* (see Chirikov (1987), p. 145), the apparatus made for Galton by Tisley and Spiller in 1873 to generate the normal distribution (see Stigler (1986)).

We must clearly be careful here not to become overawed by the new language. After all, 'stochastic' behaviour includes not only the behaviour of systems influenced throughout time by noise, but that of systems referred to in the above quotation which are merely sensitively dependent on *unknown initial conditions*. It is the fact that these conditions cannot in so many cases be known in practice (in the case of quantum mechanics this is believed to be inevitable because of the uncertainty principle) that the ideas of probability and chance are still relevant even when the system under study is regarded in principle as deterministic. Laplace's famous dictum on the complete predictability of deterministic systems becomes null and void; for chaotic systems the only kind of predictability must (e.g. in meteorology) be formulated as 'robust prediction', when found to be possible, for an appropriate set of neighbouring initial conditions.

It would, however, be a pity if the recognition today of the ubiquitous occurrence of chaos in deterministic non-linear systems is not followed in due course by a much more comprehensive classification and study of the possible structures and how best to handle them. The 'complete chaos' with which we are already familiar as 'complete randomness' is not the only kind that may appear; the strange attractors that have been already recognized need to be studied further not only in themselves but (as I remarked in Section 2 for the epidemiological models) when mixed with other very relevant stochastic effects that occur in the real world.

REFERENCES

Anderson, R. M., Grenfell, B. T. and May, R. M. (1984) Oscillatory fluctuations in the incidence of infectious disease and the impact of vaccination: time series analysis. *J. Hyg. Camb.*, **93**, 587–608.
Anderson, R. M. and May, R. M. (1983) Vaccination against rubella and measles: quantitative investigation of different policies. *J. Hyg. Camb.*, **90**, 259–325.
Bailey, N. T. J. (1957) *The Mathematical Theory of Epidemics*, p. 8. London: Griffin.
Bartlett, M. S. (1940) The present position of mathematical statistics. *J. R. Statist. Soc.*, **103**, 1–29.
—— (1949) Some evolutionary stochastic processes. *J. R. Statist. Soc.* B, **11**, 211–229.
—— (1956) Deterministic and stochastic models for recurrent epidemics. *Proc. 3rd Berkeley Symp. Mathematics, Statistics and Probability*, vol. 4, pp. 81–109. Berkeley: University of California Press.
—— (1957) Measles periodicity and community size. *J. R. Statist. Soc.* A, **120**, 48–60.
—— (1960) *Stochastic Population Models in Ecology and Epidemiology*. London: Methuen.
Berry, M. V., Percival, I. C. and Weiss, N. O. (eds) (1987) Discussion on Dynamical chaos. *Proc. R. Soc.* A, **413**, 1–199.
Chirikov, B. V. (1987) Particle confinement and adiabatic invariance. *Proc. R. Soc.* A, **413**, 145–146.
Degn, H. (1982) Discrete chaos is reversed random walk. *Phys. Rev.* A, **26**, 711–712.
Dietz, K. (1976) The incidence of infectious diseases under the influence of seasonal fluctuations. *Lect. Notes Biomath.*, **11**, 1–15.
Dietz, K. and Schenzle, D. (1985) Mathematical models for infectious disease statistics. In *A Celebration of Statistics*, pp. 167–204. New York: Springer.
En'ko, P. D. (1889) The epidemic course of some infectious diseases (in Russian) *Vrac'*, **10**, 1008–1010, 1039–1042, 1061–1063.
Fisher, R. A. and Yates, F. (1938) *Statistical Tables for Biological, Agricultural and Medical Research*. Edinburgh: Oliver and Boyd.
Gleick, J. (1987) *Chaos: Making a New Science*. New York: Viking.
Hamer, W. H. (1906) Epidemic disease in England. *Lancet*, i, 733–739.
Lewontin, R. C. (1969) The meaning of stability. *Brookhvn Symp. Biol.*, **22**, 13–24.

Martini, E. (1921) *Berechnungen und Beobachtungen zur Epidemiologie und Bekämpfung der Malaria*. Hamburg: Gente.

May, R. M. (1976) Simple mathematical models with very complicated dynamics. *Nature*, **261**, 459–467.

—— (1987) Chaos and the dynamics of biological populations. *Proc. R. Soc.* A, **413**, 27–44.

Moran, P. A. P. (1950) Some remarks on animal population dynamics. *Biometrics*, **6**, 250–258.

Olsen, L. F., Truty, G. L. and Schaffer, W. M. (1988) Oscillations and chaos in epidemics: a non-linear dynamic study of six childhood diseases in Copenhagen, Denmark. *Theor. Popln. Biol.*, **33**, 344–370.

Poincaré, H. (1896) *Calcul des Probabilités*. Paris: Gauthier-Villars.

—— (1899) *Les Méthodes Nouvelles de la Mécanique Celeste*, vols 1–3. Paris: Gauthier-Villars.

Schaffer, W. M. (1985) Can nonlinear dynamics elucidate mechanisms in ecology and epidemiology? *IMA J. Math. Appl. Med. Biol.*, **2**, 221–252.

Soper, H. E. (1929) Interpretation of periodicity in disease prevalence. *J. R. Statist. Soc.*, **92**, 34–73.

Stewart, I. (1989) *Does God Play Dice? The Mathematics of Chaos*. Oxford: Basil Blackwell.

Stigler, S. M. (1986) *The History of Statistics*, p. 277. Cambridge: Harvard University Press.

Takens, F. (1981) Detecting strange attractors in turbulence. In *Dynamical Systems and Turbulence*, (eds D. A. Rand and L. S. Young), pp. 366–381. New York: Springer.

Thompson, J. M. T. and Stewart, H. B. (1986) *Nonlinear Dynamics and Chaos*. New York: Wiley.

Wright, S. (1968–78) *Evolution and the Genetics of Populations*, vols 1–4. Chicago: University of Chicago Press.

Zeeman, C. (1988) Stability of dynamical systems. *Nonlinearity*, **1**, 115–155.

DISCUSSION OF THE PAPER BY BARTLETT

Professor Howell Tong (University of Kent, Canterbury): This is a truly remarkable recurrent event. How many of us were here half a century ago when tonight's speaker gave his other equally far-sighted paper? That was really before my time.

Tonight's paper is about chaotic time series. A simple and yet effective way to depict such a time series is a *phase diagram*. In discrete parameter time series, the idea goes back to at least Yule (1927). I have been advocating the use of a directed scatterplot in non-linear time series analysis for at least 10 years. Fig. 2 is such an illustration of a possibly chaotic time series taken from Tong (1990), which also lists several references on chaos. Similar plots were independently introduced by, principally, Takens in about 1980 and have been extensively used in chaos studies (see Schaffer (1985)).

Now, visualize the *stretching* and the *folding* actions of a baker or a noodle maker. The former induces the sensitivity to initial values and the latter ensures global boundedness. These two ingredients are equally essential for the generation of chaos. Moreover, the *self-similarity* of the noodle sticks is one of the many hallmarks of chaos. Other hallmarks include such exotic concepts as *fractional dimension*, *period doubling*, *bifurcation*, etc. Professor Bartlett has quite rightly urged us to try to understand chaos and

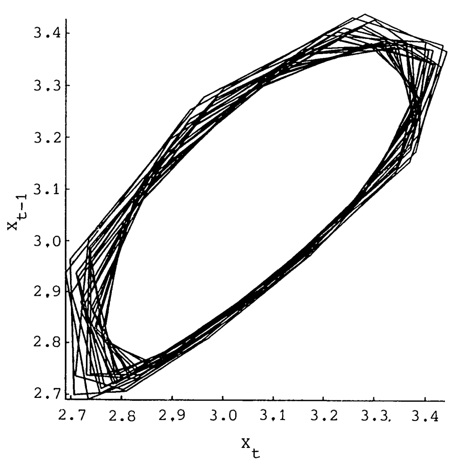

Fig. 2

emphasized its intimate relationship with chance. I would like to remind Fellows that the English word CHANCE came from the old French word CHEANCE, which in turn came from the late Latin word CADENTIA, meaning the falling (or perhaps the dropping) of a penny. A statistician is thus in daily contact with chaos. Fig. 2 of DeGroot (1986) highlights the intimate connection between our main randomization mechanism and chaos. What potential benefits can we, statisticians, reap from a study of this fascinating new concept? I list some below.

(a) We can calibrate time series by their improvements in long-term predictive performance over a linear analysis.
(b) We can exploit noise reduction algorithms developed by the dynamicists. (Incidentally, I was told that calling anybody a chaoticist would be taken as an insult.)
(c) We can borrow some of their graphical techniques, such as the Poincaré section.

What can we contribute in return? Here are some suggestions.

(i) A dynamicist would typically start with the deterministic model

$$X_t = f(X_{t-1}, \ldots, X_{t-p}). \tag{7}$$

However, X_t may never be observed exactly. Instead, we observe, say,

$$Y_t = X_t + \epsilon_t, \tag{8}$$

where ϵ_t are independent and identically distributed random variables. In this case

$$Y_t = f(Y_{t-1} - \epsilon_{t-1}, \ldots, Y_{t-p} - \epsilon_{t-p}) + \epsilon_t, \tag{9}$$

a non-linear autoregressive moving average ARMA(p, p) model. Given $Y_1, \ldots, Y_n$ how can we determine p *if f is unspecified*? Here p is related to the so-called *embedding dimension*. Consider a simpler model, namely

$$Y_t = f(Y_{t-1}, \ldots, Y_{t-p}) + \epsilon_t. \tag{10}$$

Elsewhere (Tong, 1990) I have suggested three methods.

Let $p \ll K \ll N$ be fixed. Let $\delta_N(\cdot)$ denote the usual kernel in nonparametric regression. Define f_t: $R^p \to R$ by

$$f_t(z_1, \ldots, z_p) = \sum_{s=K}^{t-1} y_s \left\{ \prod_{i=1}^{p} \delta_N(z_i - y_{s-i}) \right\}^{1/p} \Big/ \sum_{s=K}^{t-1} \left\{ \prod_{i=1}^{p} \delta_N(z_i - y_{s-i}) \right\}^{1/p}.$$

Define h_j: $R^{2p} \to R$ by

$$h_j(z_1, \ldots, z_p; z_{-1}, \ldots, z_{-p}) = \frac{\sum_{s=K}^{N-K} y_s \prod_{i=1}^{p} \{\delta_N(z_i - y_{s+i})\, \delta_N(z_{-i} - y_{s-i})\}^{1/(2p-1)}}{\sum_{s=K}^{N-K} \prod_{i=1}^{p} \{\delta_N(z_i - y_{s+i})\, \delta_N(z_{-i} - y_{s-i})\}^{1/(2p-1)}},$$

where all summands which contain y_j are excluded. Method (a) is the predictive residual approach:

$$\min_p \left[\sum_{t=K}^{N} \{y_t - f_t(y_{t-1}, \ldots, y_{t-p})\}^2 \right].$$

Method (b) is the Akaike information criterion approach:

$$\min_p \left[\sum_{t=K}^{N} \{y_t - f_{N+1}(y_{t-1}, \ldots, y_{t-p})\}^2 \right].$$

Method (c) is the cross-validation approach:

$$\min_p \left[\sum_{t=K}^{N-K} \{y_t - h_t(y_{t+1}, \ldots, y_{t+p}; y_{t-1}, \ldots, y_{t-p})\}^2 \right].$$

(Method (b) might be related to Tjøstheim's unpublished approach.)

(ii) How large a sample do we need to obtain reliable estimates of the Lyapunov exponents, which quantify local instability, and the dimensions of the attractors?

(iii) Following (i), suppose that we write more suggestively (see Tong (1990), pp. 127 and 150)

$$\begin{aligned}\text{real time series} &= \text{trend} + \text{irregularity}\\ &= \text{skeleton} + \text{stochastic component}\\ &= \text{low dimensional attractor} + \text{high dimensional attractor}\\ &= \text{Yin} + \text{Yang},\end{aligned}$$

associating stochastic noise with the second summands. Given N observations, how do we test that the low dimensional attractor is a strange attractor, a limit cycle or a limit point? This problem seems solvable within the parametric framework (Tong, 1990).

I have one question and one specific comment. The SEIR model (1) is an example of the so-called bilinear models in control engineering. Brockett (1977) has shown that such models (keeping m, a, b and g constant) cannot admit limit cycles. Can they admit chaos? Whittle (1963) has discussed a similar curious example to that in Section 4.

Finally, in the words of Hesiod: 'First verily was created *chaos* (χάος), and then broad-bosom Earth'. I am sure that out of our involvement in chaos something beautiful will emerge. It is my great pleasure to be given the privilege of proposing a warm vote of thanks to Professor Bartlett for a most timely paper. I wish him many happy returns.

R. L. Smith (University of Surrey, Guildford): The subject of 'chaos' is one which has gained a great deal of both popular and scientific acclaim, yet it is one in which, up to now, statisticians have not played a significant role. It is surely necessary that they should, and we should all be grateful to Professor Bartlett for introducing this subject to the Society, as well as for giving us some fascinating insights of his own.

In my own remarks I would like to focus on the statistical problem of inferring chaos from a time series. Three elements of chaos theory seem particularly relevant to this: the Takens embedding theorem, mentioned by Professor Bartlett in Section 3 of his paper, the concept of a strange attractor and its associated dimension, and Lyapunov numbers.

The Takens theorem states, loosely, that if $\{x_n = s(n\tau),\ n = 1, 2, \ldots\}$ is a discrete-time scalar cross-section from a deterministic continuous time chaotic process with an I-dimensional strange attractor, then there exists an $m \leqslant 2I+1$ and a smooth function f such that

$$x_n = f(x_{n-m}, x_{n-m+1}, \ldots, x_{n-1}). \tag{11}$$

The statistical consequence of this is that it reduces the problem of detecting chaos to one of finding mappings of the form (11) (presumably, with some allowance for random noise) in discrete-time series.

Given such a mapping, its *attractor* is the subset of $(m+1)$-dimensional space in which the vectors $(x_{n-m}, \ldots, x_n)$ eventually lie. In a chaotic system, this will typically be a *strange attractor* of fractional dimension. Consequently, an important aspect of detecting chaos is the ability to recognize fractional dimension.

Lyapunov numbers are a measure of the instability of a chaotic system. The example in Professor Bartlett's Section 4 illustrates this in a clear, if elementary, fashion. This is a function $x_{n+1} = f(x_n)$ for which $f'(x)$ is almost everywhere 3. Consequently, if $\{x_n\}$ and $\{y_n\}$ are two trajectories with $|x_0 - y_0| = \epsilon$, we expect $|x_1 - y_1| = 3\epsilon$, $|x_2 - y_2| = 3^2\epsilon$, and so on until the two trajectories lie in different thirds of the unit interval. The time taken to achieve this is of the order $n = -\log \epsilon / \log 3$, and it can be checked that Professor Bartlett's numerical examples (with $\epsilon = 10^{-6}$ and $\epsilon = 10^{-3}$) are consistent with this. For this system, 3 is the *Lyapunov number*. For an m-dimensional system there are m Lyapunov numbers, defined as eigenvalues of a limiting Jacobian matrix, and chaos is characterized by at least one of the Lyapunov numbers being greater than 1.

The most obvious statistical problem is that of estimating dimension of an attractor. The first point to note is that there are many conflicting definitions of fractional dimension; Farmer *et al.* (1983) review this topic, though many of their results are conjectural. The very recent monograph of Falconer (1990) goes into these questions from a much more rigorous viewpoint. Of estimation methods, *box counting* algorithms (Barnsley, 1988) are the most easily understood, but probably better overall are estimation methods based on nearest neighbour distances, of which the Grassberger and Procaccia (1983) algorithm has achieved a certain amount of fame and, if we believe Ruelle (1990), notoriety. Smith (1988) purported to show that, for an attractor of dimension d to be estimated to within 5% of its true value, a sample size of at least 42^d is required. Some calculations of my own, based on bias versus variance considerations of the kind familiar to theoretical statisticians, have suggested that this assessment is too pessimistic, but the general message of Smith and Ruelle, that excessively large sample sizes are needed to identify attractors of high dimension, seems to be correct.

There are many other statistical problems connected with the Grassberger–Procaccia algorithm and its relatives. A particularly important issue is the effect of experimental noise, since relatively modest amounts of noise can make it virtually impossible to estimate the dimension accurately regardless of sample size.

Some other references on dimension estimation are the review paper of Broomhead and Jones (1989) and two papers concerned primarily with fractals (Cutler and Dawson, 1989; Taylor and Taylor, 1991) but nevertheless relevant to the problem of detecting chaos in time series.

The estimation of Lyapunov exponents has also attracted attention (e.g. Sano and Saweda (1985) and Eckmann *et al.* (1986)), and there is a growing literature on the actual estimation of f in equation (11); e.g. Farmer and Sidorowich (1987, 1988) and Casdagli (1989). I believe that all these papers merit careful study by statisticians, and pose many problems for statistical research.

In conclusion, where Professor Bartlett has focused on the interpretation of chaotic models and in particular their relation with much longer-established ideas of stochastic modelling, I have concentrated my own remarks on aspects of statistical inference. I believe, however, that from both points of view the message is that probabilists and statisticians should be taking a serious interest in the ideas and challenges posed by chaos. Professor Bartlett has taken an important first step in this direction, and it gives me great pleasure to second the vote of thanks to him.

The vote of thanks was passed by acclamation.

Rodney C. Wolff (University of Oxford): This paper aims to raise questions. It certainly hits the target very rapidly. The first contention appears to be prompted by the initial quote from Schaffer (1985). Before that, Schaffer mentions the difficulty of quantifying noise in erratic data, where 'erratic' embraces chaotic and random effects. Noise and chaos together could result in one masking the behaviour of the other. This topic is part of the increasingly popular theme in chaos theory and statistics: to what extent are chaos and randomness disjoint entities?

Regarding the quantification problem, the correlation integral may be of use. The correlation integral was proposed by Grassberger and Procaccia (1983) as a means of obtaining the embedding dimension of the series attractor. The integral can distinguish between a sequence of independent identically distributed random variables and one generated by a deterministic non-linear difference equation. If, given the multivariate time series in Section 3 of this paper, we define vectors

$$V(t)=\{X(t), \ldots, X(t+(p-1)h)\}' \qquad (t=1, \ldots, N') \quad (N'=N-ph+h)$$

then the correlation integral is given by

$$C(r;\, p)=\lim_{N\to\infty}\left[\frac{2}{N(N-1)}\sum_{i=1}^{N'-1}\sum_{j=i+1}^{N'} I\{r-\|V(i)-V(j)\|\}\right]$$

where $I(x)=1$ if $x\geqslant 0$ and $I(x)=0$ otherwise; where $r>0$ is fixed; and where p is a fixed natural number called the phase dimension. If the series $\{X(t)\}$ is purely random with no systematic effects, it should be that $C(r;\, p)\propto r^p$, as p increases. If $\{X(t)\}$ is chaotic then saturation in the power law occurs, namely $C(r;\, p)\propto r^{p'}$ for all $p>p'$, for some constant p', the estimated embedding dimension. Wolff (1990) has shown that if $\{X(t)\}$ is generated by a simple linear Gaussian autoregressive moving average model then such a saturation eventuates, albeit extremely slowly. Thus a random process, if not somehow classifiable as chaos, can mimic chaos. This is just one example of the difficulty in distinguishing between chaos and randomness, the step which must precede quantifying the noise.

Another difficulty is that dynamical systems may amplify noise, yet on their own may be globally stable. The noise may arise from external effects or complicated phenomena, such as thermal motions. Deissler and Farmer (1990) consider this problem when the quantities of noise are small. The extension to cases of large amounts of noise would be of interest.

For noisy or random looking data, a test for randomness against an alternative of a particular type of chaotic process may be provided by the correlation integral. Brock *et al.* (1986) have used the integral in a central limit theorem to test for independence against the complementary alternative. The chaotic alternative may provide a more powerful test.

These comments illustrate that the 'still waters' of once familiar erratic data run deep . . . and chaotically, too.

T. Subba Rao and Jingsong Yuan (University of Manchester Institute of Science and Technology): The relationship between chaotic models and non-linear time series models is interesting, because all chaotic models are non-linear, but not all non-linear models are chaotic. For example, the logistic model discussed by Professor Bartlett looks like a bilinear model (see Subba Rao and Gabr (1984)), but with different properties. However, the techniques developed for analysing non-linear models may prove useful, though they are not sufficient. In this context we refer to Professor Bartlett's comment in Section 3 that conventional spectral analysis is not adequate for analysing chaotic models; a similar observation has been made by Brock (1986) and Sakai and Tokumaru (1980). This is because the second-order properties for some chaotic models are similar to linear models. This is quite common in non-linear models. For example, Subba Rao (1981) has shown that some bilinear models have second-order properties similar to autoregressive (AR) moving average models and, therefore, for bilinear models, Sesay and Subba Rao (1988) have developed difference equations for higher order cumulants for the identification of these bilinear models. To illustrate these ideas, we consider an alternative chaotic model, called a tent map, which is given by

$$X_{t+1} = \begin{cases} X_t/a & \text{if } 0 \leqslant X_t \leqslant a, \\ (1-X_t)/(1-a) & \text{if } a \leqslant X_t \leqslant 1. \end{cases}$$

Sakai and Tokumaru (1980) have shown that the second-order covariances of the tent map are similar to the linear AR(1) model $X_{t+1} = (2a-1)X_t + e_{t+1}$, and therefore the conventional second-order spectral analysis cannot play a useful role in the analysis of non-linear (chaos) models. This suggests that higher order spectra, such as bispectra, may be useful. It is well known (see Subba Rao and Gabr (1984)) that if the process is linear the ratio $B(\omega_1, \omega_2) = |f(\omega_1, \omega_2)|^2 / f(\omega_1) f(\omega_2) f(\omega_1 + \omega_2)$, where $f(\omega_1, \omega_2)$ is the bispectrum of the process and $f(\omega)$ is the spectrum of the process, is constant for all ω_1 and ω_2 (the bispectrum is zero if the process is Gaussian). A time series of length 256 is generated from this model with $a = 0.25$ and $X_1 = 0.6$; the plot of the data is given in Fig. 3; the estimated $B(\omega_1, \omega_2)$ is given in Fig. 4. (The details of the estimation procedure are given in Subba Rao and Gabr (1984).) Though the shape of the spectrum is similar to that of an AR(1) model the shape of $B(\omega_1, \omega_2)$ rules out the possibility that the process is linear. The presence of the ridge along the line $\omega_1 + \omega_2 = 1.2\pi$ is interesting, and we do not yet know how to interpret this in relation to chaotic dynamics of the model. In view of this preliminary analysis, we believe that the higher order spectral analysis of chaos models may play an important role in future studies.

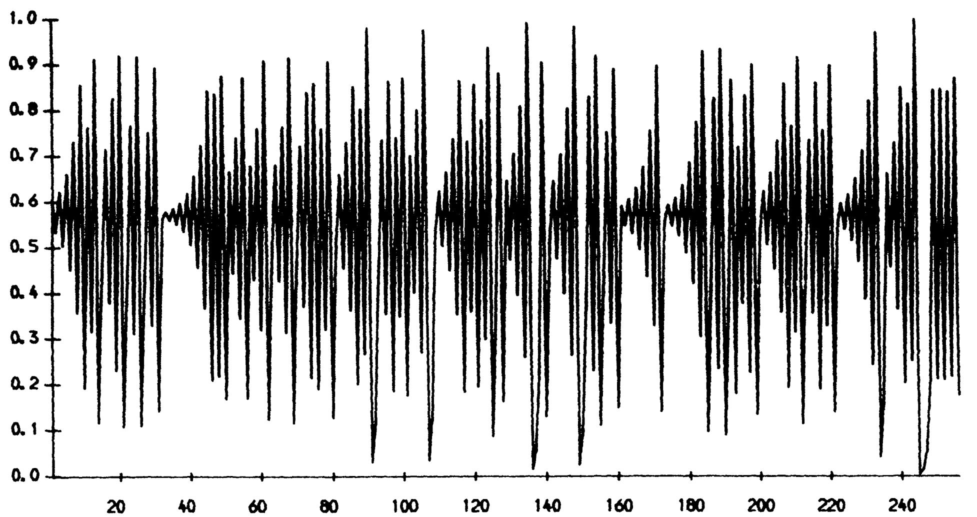

Fig. 3. Chaos data with $A = 0.25$ and $X(0) = 0.6$

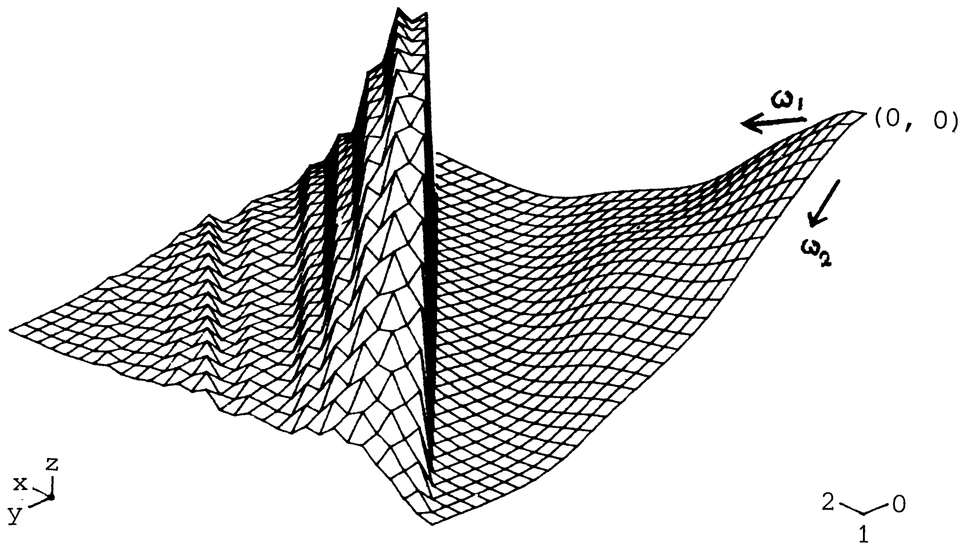

Fig. 4. Estimated modulus of a bispectrum (normalized)

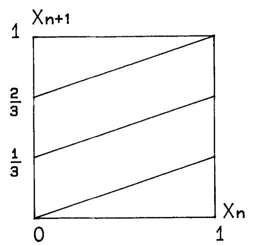

Fig. 5. Scatterplot of X_{n+1} against X_n for the uniformly distributed first-order linear autoregressive model when $k=3$

A. J. Lawrance (University of Birmingham): This delightful paper raises many interesting questions, and is also delightfully brief. I wish to make a couple of hopefully useful comments on Professor Bartlett's 'curious example' of a first-order uniformly distributed linear autoregressive time series model which is deterministic in reversed time. His memory concerning Professor Murray Rosenblatt is likely to be correct in that I know of four of his works mentioning the 'curious example' (Rosenblatt, 1979, 1980, 1985, 1986); the 1986 paper is associated with the first two. To quote from the 1980 paper, 'The process is purely non-deterministic going forwards in time and purely deterministic going backwards in time'. Another instance of this change in model on reversing time is given by the exponential linear autoregressive model (4), studied by Gaver and Lewis (1980), which in reversed time is the non-linear autoregressive model $X_{n+1}=\min\{X_n/\rho,\ E_n/(1-\rho)\}$, where E_n is an exponential innovation; this was shown by Chernick *et al.* (1988). Regarding Professor Bartlett's uniform model, one accessible approach for the non-specialist is via moment-generating functions; we find $\phi_Y(t)=\phi_x(t)/\phi_x(\rho t)=\rho(e^t-1)/(e^{\rho t}-1)$ which appears from factorization to be proper only when $\rho=\pm 1/k$, for integer $k\geqslant 2$, and then Y has a 'curious' discrete uniform distribution on $(0, 1, \ldots, k-1)/k$. The reversed deterministic aspect is most easily seen from plotting X_{n+1} against X_n, as illustrated for $k=3$ in Fig. 5; only realizations along the sloping lines are possible, whence the deterministic backward behaviour is evident, e.g. $X_n=3X_{n+1}-2$ when $\frac{2}{3}<X_{n+1}\leqslant 1$. This deterministic behaviour, given fully by equation (6), exhibits a beguilingly simple sensitivity to initial conditions. The sensitivity of one-step prediction in the backward sense given X_n

near $(1, 2, \ldots, k-1)/k$ is a further curiosity of this model. The negatively correlated case of this model is similarly curious, and is easily seen from its structure as

$$X_{n+1} = \frac{1}{k}(1 - X_n) + Y_{n+1},$$

and is thus also deterministic in the backward sense. These are minor concerns in the infinite world of chaos, but my broader view from this paper is that chance and chaos should treat each other as long lost friends.

Dr Eric Renshaw (University of Edinburgh): In his last paragraph Professor Bartlett remarks that 'strange attractors . . . need to be studied further not only in themselves but . . . when mixed with other very relevant stochastic effects that occur in the real world'. To develop this point let us take a brief foray into the field of non-linear models (Renshaw, 1990).

The Lotka–Volterra process, e.g.

$$\begin{aligned} \mathrm{d}N_1/\mathrm{d}t &= N_1(1.5 - 0.1N_2) \qquad &\text{(prey)},\\ \mathrm{d}N_2/\mathrm{d}t &= N_2(-0.25 + 0.01N_1) \qquad &\text{(predators)}, \end{aligned}$$

was one of the first non-linear models to excite population biologists since it gives rise to a family of closed trajectories. Such was the appeal of this result that biologists spent several unsuccessful years trying to mimic persistent predator–prey cycles in the laboratory. The difficulties involved are exposed by a stochastic analysis: simulations starting from either the equilibrium point (25, 15) or an extreme position such as (3, 2) almost always lead to extinction of one of the species before the end of the first cycle.

That neutrally stable deterministic cycles give rise to divergent stochastic cycles suggests that convergent deterministic cycles might lead to sustained stochastic oscillations. Not only does this assertion hold true when a carrying capacity is included in the prey birth component, but the qualitative time series properties of this modified stochastic model are in accord with the original deterministic Lotka–Volterra prediction! Agreement between both deterministic and stochastic behaviour can, however, be obtained by considering models with stable deterministic cycles, such as the Holling–Tanner model with its non-linear predator response.

Let us now consider chaos through the (unscaled) equation (3), i.e.

$$N_{t+1} = N_t\{(1+r) - rN_t\}.$$

This discrete-time function successively generates

(a) damped oscillations $(0 < r < 2)$,
(b) stable two-point cycles $(2 < r < 2.5)$ and
(c) stable 2^2, 2^3, 2^4 . . . -point cycles $(2.5 < r < 2.57)$,

before exhibiting

(d) three-point quasi-cycles for r just above 2.57.

This transition into a regime of apparent chaos may be relevant to temperate insect populations. Though stochastic behaviour shows (a) random meandering about the equilibrium point, (b) two-point cycles and (d) three-point cycles, the delicate structure (c) collapses to a basic two-point cycle. Spectral analysis shows that the power associated with cycles of order higher than 2 is so small that it becomes swamped by the addition of stochastic effects. So complex chaotic regimes in which, for example, '93 different intrinsic 11-point cycles each cascade down through their harmonics of stable cycles with periods 11×2^n' may be of little practical interest.

The extent to which non-linear deterministic systems retain their underlying behaviour when stochastic effects are introduced clearly needs considerably more understanding than exists at present if such systems are to gain a foothold in established statistical practice.

Mr A. J. Mayne (Milton Keynes): I would like to make a few further remarks about the sort of model which Dr Renshaw discussed. It had also occurred to me that it would be worthwhile to investigate the result of putting some stochasticity into Robert May's models, not random noise on the right-hand side of equation (3), but instead making the parameter k itself a random variable. It would be interesting to know what would happen if that were done. Are there similar disturbances to the deterministic periodicities and chaotic phenomena which occur in the other case? As far as I know, nobody has yet tried this.

Another interesting topic on the subject of iteration would be the sensitivity of the iteration to small disturbances in the functional form. Again considering equation (3), suppose that we add, first, in a deterministic model, a relatively small additional term, let us say a term multiplied by a certain parameter c, which might be a small parameter compared with k. How does the phenomenon change as the value of c changes? It would also be interesting to investigate the combination of that with stochasticity.

Turning to the end of Professor Bartlett's paper, I would like to make a few observations on the nature of uncertainty. He has mentioned stochasticity or randomness and chaos as two different types of uncertainty. Also, with pseudorandomness perhaps being viewed as a special case of chaos, maybe randomness itself can be viewed as a special case of chaos.

However, there is another type of uncertainty which is found quite frequently in the literature, called fuzziness. What happens when all three of these operate together? I suggest that this is a further area of research.

We have already heard about the initial conditions, which again Professor Bartlett mentions at the end of his paper. We know how sensitive subsequent phenomena can be to initial conditions. Are there circumstances in which what I would call 'intermediate conditions' can also have a significant effect on the phenomena? An example of intermediate conditions would be stochastic variations as the iterations of the phenomena unfold. It would be quite interesting to look into that.

Finally, I raise the question whether the concepts of chaos give us any ground for further generalizations of the concepts of statistical distribution functions and statistical distribution functionals. That would be worth looking into.

Mr A. R. Thatcher (New Malden): The curious example in Section 4 describes a process which is stochastic when time flows forwards but is deterministic when the flow of time is reversed. Many physicists are concerned about a problem which is described as 'the arrow of time'. The main laws of physics are unchanged if one changes the sign of t, so what is it which determines the direction in which things happen?

Prigogine and Stengers (1986) link this question to information theory and the theory of chaos. Because chaotic situations can occur even in a deterministic universe, reversing the direction of time would involve transfers of infinite amounts of information.

One wonders whether Professor Bartlett's example may be relevant to this problem. It certainly gives an amazingly neat illustration of a connection between chance, determinism and the direction of time.

Professor Tohru Ozaki (Institute of Statistical Mathematics, Tokyo): On the whole, I agree with Professor Bartlett's critical view of the chaos modelling approach in Sections 2 and 5, although his criticism sounds almost too modest compared with some comments of chaoticists (as we might call scientists who believe in determinism) on stochastic modelling.

The model which chaoticists suggest is a special case of the following non-linear state space model:

$$\dot{\mathbf{z}} = \mathbf{f}(\mathbf{z}|\boldsymbol{\theta}) + \mathbf{v}(t)$$
$$\mathbf{x}_t = \mathbf{z}_t + \mathbf{w}_t$$

where $\mathbf{v}(t)$ is a continuous time system noise and $\mathbf{w}_t$ is a discrete time observation noise. Chaoticists claim that $\mathbf{v}(t) = 0$ and $\mathbf{w}(t) = 0$. However, as we cannot deny the uncertainty principle in quantum mechanics, we cannot assume $\mathbf{w}_t = 0$. If $\mathbf{v}(t)$ really was zero, then the maximum log-likelihood of the model should occur when the estimated variance–covariance matrix Σ_v of $\mathbf{v}(t)$ is close to zero. We know that this never happens in real data analysis and the maximum likelihood estimate of Σ_v is disappointingly large. Thus the main concern of applied time series analysts is to find the $\mathbf{f}(\mathbf{z}|\boldsymbol{\theta})$ of the least possible Σ_v.

Non-linear modelling and the determinism of chaoticists are different things. I agree that chaoticists have introduced some interesting $\mathbf{f}(\mathbf{z}|\boldsymbol{\theta})$ which might be useful in reducing the variance of $\mathbf{v}(t)$ (and increasing the likelihood, which implies a smaller Kullback–Leibler distance of the model). However, I am not so optimistic to expect the variance to be zero in real data analysis. Some of the chaoticists' models, especially discrete-time models such as equation (3), look obviously useless in data analysis, because with these models chaotic trajectory is produced only in a narrow region of parameter space. When we model stochastic phenomena, we expect that two models with similar parameters would exhibit similar features. This cannot be expected with these discrete-time chaos models. If we have to estimate unknown model parameters and the initial value and if the future behaviour is so sensitive to a microscopic deviation of the unknown initial value, there is not much point to replacing stochastic prediction by deterministic prediction.

Chaoticists' determinism is a revival of the ghost of Laplace's daemon which resisted the surge of Maxwell's and Boltzmann's ideas of statistical mechanics so strongly at the end of the last century. As Professor Bartlett says, we statisticians should not be overawed by new language and should continue the stochastic approach not only with conventional models but also mixed with some of the chaoticists' non-linear models.

Professor J. Durbin (National University of Singapore, Kent Ridge): I wish to add my congratulations to Professor Bartlett on his golden jubilee paper and I endorse his view that the development of chaos theory is important for statistics. Many of us have felt uneasy from time to time about the transition from deterministic phenomena such as coin tossing to the chance mechanisms that we employ to analyse them. The notion of chaos makes the transition from a deterministic to a probabilistic description more intelligible. Subjectivists argue that probability arises only from uncertainty in the human mind and should consequently be modelled by a subjective theory. However, work on chaos has established that objective probability models for large-scale phenomena have just as much validity as other mathematical models in physics.

I hope that the day will come when virtually all statisticians accept both objective and subjective probability models as valid, the choice between them being made on each specific occasion on pragmatic grounds, rather than decided *a priori* on ideological grounds. Developments in chaos theory should help to speed the dawn of that happy day.

These matters are discussed at greater length in my Presidential Address to the Society (Durbin, 1987), particularly Section 4.2.

K. Dietz and D. Schenzle (Institute of Medical Biometry, Tübingen University): The epidemiological model under consideration is the SEIR model (1) with a periodic contact rate of the form

$$b(t) = b_0\{1 + b_1 \cos(2\pi t)\}.$$

What Schaffer (1985) claims is that 'under certain conditions' this model yields a 'chaotic' phase portrait which is 'in excellent agreement' with phase portraits reconstructed from actual measles incidence data. However, for the following reasons we share Professor Bartlett's more sceptical view about the validity and relevance of 'chaos' in connection with infection incidence patterns.

(a) According to Schaffer's choice ($b_0 = 1800$/year) half of all children should have had measles by age 2.3 years. This is in conflict with the observed median age at infection of about 5 years.
(b) Model (1) cannot reproduce the empirically observed characteristic difference in the variation of the contact rate during years with high and low measles incidence (London and Yorke, 1973; Fine and Clarkson, 1982).
(c) For a more rigid test of the model it is necessary to compare model results with detailed measles incidence data throughout the year. Moreover, how does model (1) with Schaffer's parameter values perform when simulating the effect of mass vaccination against measles?
(d) Finally, we would like to know the minimum prevalences for the predicted chaotic incidence patterns. Are these minimum prevalences still sufficiently high to justify a deterministic model at all?

In our opinion too there is no need for a dramatic switch in theoretical thinking and epidemiological modelling. The limited applicability of the simple global mass action model (1) is by now well known. Especially for the so-called childhood infections there exist more detailed deterministic models, the parameters of which have an epidemiological interpretation (e.g. Schenzle (1984)). These models may be modified to allow for chance effects, which we consider a natural element in infection transmission.

The following contributions were received in writing after the meeting.

A. Babloyantz (Free University of Brussels): In this paper Professor Bartlett points to the importance of the relatively new field of deterministic chaos in the field of epidemiology. I wish to comment on the existence of deterministic chaos in the electrical activity of the cortex.

The global electrical activity of the human brain, the so-called electroencephalogram, can be measured non-invasively from the scalp. From the measurement of such a single lead one constructs a multidimensional phase space following Takens theorem as stated in this paper. Once an approximate topological equivalent to the original attractor is found the nature of the underlying dynamics can be assessed by various techniques. For example one measures the correlation dimensions, the Lyapunov exponents and metric entropies. Our group and later other research groups found that β-waves do not show any marked coherence, however, as the α-waves set in the cerebral activity is governed by deterministic chaos.

As the sleep cycle starts the coherence of cerebral activity increases and the dimension of the chaotic attractors decreases. In normal subjects the lowest dimension of chaotic attractors is seen in sleep stage four. In rapid eye movement sleep the brain dynamics switch back to a much less coherent state. In severe pathologies such as the Creutzfeld–Jakob disease and the '*petit mal*' seizure the chaotic attractors are of a very low dimension; only four or three degrees of freedom are needed for a description of the system's dynamics.

In animal studies, such a dynamical approach may give useful information about the differences in the activity of various layers of the cortex.

Sir David Cox (Nuffield College, Oxford): Professor Bartlett's paper raises important issues about a fascinating topic. There are several quite distinct aspects with a statistical flavour including the following.

(a) In those situations in which a simple controlling mechanism may operate, how do we test agreement with a postulated model and with a postulated dimension and unspecified model, incorporating preferably some allowance for a stochastic component, distinguishing between an error of observation and a random element incorporated into the process?
(b) What are the statistical properties of chaotic processes? Something is known about spectra and stationary distributions.
(c) Are some of the data analysis techniques used in chaos studies, e.g. the Grassberger–Procaccia algorithm, useful for other statistical purposes? There is some connection with the extensive current work on nonparametric regression.
(d) Is the usefulness of empirical prediction methods based on ideas from chaos theory confined to those situations in which we can reasonably expect one simple controlling mechanism to operate with at most minor 'noise'? (That would rule out applications in many fields!) Again there are connections with nonparametric density estimation. For an excellent discussion, see Farmer and Sidorowich (1987).

The chaotic processes in the literature resemble stochastic processes with continuous sample paths. It is reasonable to ask whether there are chaotic analogues of point processes. Walter Smith and I a few years ago (Cox and Smith, 1953) studied what might be regarded as one such, a hypothesis suggested by Professor B. Katz that some spike train data might be generated by superimposing a fairly small number of strictly periodic signals. Such a deterministic system would indeed generate data looking very like a Poisson process, locally at least.

Professor A. P. Dawid (University College London): Chaos theory presents a philosophical challenge for traditional probability theory. The phenomena it purports to describe—in particular, short-term predictability combined with long-term unpredictability—have not traditionally been regarded as suitable candidates for probabilistic explication. Nevertheless, there are many similarities with probabilistic concepts of randomness, and it is interesting to ask whether there might be some broader view which encompasses both probability and chaos.

It should be possible to develop such a general theory on the basis of the algorithmic approach to randomness, of which there are several variants—see for example Kolmogorov and Uspenskii (1987). In this approach, one attempts to quantify the non-randomness of a data sequence by the extent to which it can be explained by means of a suitably circumscribed collection of deterministic descriptions. The basic approach uses finite algorithms as descriptions. A sequence may then be termed 'chaotic' if any successful algorithm is essentially as long as the sequence itself. We can go on to consider the extent to which a data sequence can be explained using a broader class of descriptions—e.g. optimal codes based on probability distributions (Rissanen, 1989). The upshot of these considerations is an extension of traditional probability theory, which allows both deterministic and non-deterministic explanations, as well as various degrees of approximate randomness (Vovk, 1987).

In my own approach (Dawid, 1985), I considered the ability of a data sequence to support description by one-step-ahead probability forecasts, using an information base consisting of full or partial information about previous values. It then turns out that a sequence can be deterministic for one information base but probabilistic for another, coarser, one. It even appears possible for a deterministic sequence to transform into a stochastic one on introducing any inaccuracy, however small, into the information base. Although this approach is based on short-term forecasting, it would seem to chime well with the emphasis of chaos theory on long-term uncertainty as a feature of extreme sensitivity to initial conditions.

F. Drepper (Jülich Research Centre): Starting from meteorology and physics in the 1960s and 1970s the idea that chaotic low dimensional dynamical systems are not only useful as pseudorandom number generators but also relevant as deterministic models of fluctuating real systems is now penetrating the empirical side of the life sciences. The special importance of Bartlett's paper as part of this historic process lies in the emphasis put on 'the assessment of the role of chaos in conjunction with other factors' of unpredictability. Although there is an extensive literature on the—indeed—fascinating properties of purely deterministic dynamical systems, there is relatively little known about the subtle interactions between random and (low dimensional) non-linear sources of unpredictability—in particular about methods to distinguish them in empirical time series.

In this context I would like to draw attention to new methods of time series analysis based on the scaling behaviour of different measures of information. The best known example of these is the Grassberger–Procaccia method, which has been discredited partially because of its misuse and partially because of its weakness when dealing with additional noise sources and with metastable transients.

An alternative to the Grassberger–Procaccia method has been developed (Drepper, 1988; Drepper *et al.*, 1990). It is based on the information content of *conditional* one-step predictions and seems to be better suited to disentangle the role of chance and low dimensional unstable deterministic dynamics in generating the unpredictability in the behaviour of individual and collective biological systems. The specific feature of deterministic motion lies in that for these systems the uncertainty of a conditional prediction becomes independent of its accuracy, if the precision of the prediction is coupled to the precision of the knowledge of the past. Empirically obtained as well as model-based time series of measles incidence patterns have been used as examples to demonstrate the usefulness of an 'unpredictability profile' as a new instrument of non-linear time series analysis.

Mr D. A. Elston and Dr C. A. Glasbey (Scottish Agricultural Statistics Service, Edinburgh): Professor Bartlett observes that pseudorandom numbers are chaotic, rather than truly random, and hence that randomness can be thought of as an extreme form of chaos. Indeed, every random process has a chaotic process which is indistinguishable from it in any finite realization (simply replace every random variate by a pseudorandom variate). Is the converse of this statement also true, or are there forms of chaos which can be distinguished from random processes? The answer may lie in the development of theoretical models and analysis of observations made at suitably small time intervals. For example, the long run distribution is beta$(\frac{1}{2}, \frac{1}{2})$ for values generated by the chaotic difference equation

$$x_{n+1} = 4x_n(1-x_n)$$

(taking $k=4$ in equation (3)). If we were to observe the system infrequently, the xs would appear to be randomly distributed whereas a plot of x_{n+1} against x_n would reveal the quadratic relationship. We wonder whether a chaotic process will eventually supplant randomness at the heart of quantum physics.

Some statistical methods have already been used in the study of chaotic systems. Broomhead and King (1986) used principal components analysis to find the embedding dimensions of strange attractors. Ramsey and Yuan (1990) performed a simulation study to investigate the sampling properties of an estimator of the fractal dimension of strange attractors. Farmer and Sidorowich (1987) based forecasts of chaotic systems for which time series data were available on past instances in the data which were most similar to the current state. This leads to forecasts which are robust in that they do not require estimation of an underlying model. Artificial neural networks have also been used to implement this form of forecasting.

Of the many statistical problems in the study of chaotic systems, two which draw immediate attention are the following.

(a) Find goodness-of-fit tests for models which generate fractal structures. Although models may agree qualitatively with observation in leading to fractals of the correct dimension, what quantitative methods can be used? Under what circumstances can data be said to support the existence of a particular strange attractor predicted by a theoretical model?

(b) Describe the joint distribution of variables constrained to lie on a strange attractor in multivariate space, thus generalizing the above result for the beta distribution. Of the many possible applications, the following springs immediately to mind: estimate extreme values of meteorological variables when changing patterns of climate make reliance on historical records misleading.

Dr R. W. Farebrother (University of Manchester): I should like to congratulate Professor Bartlett on a most stimulating paper. I will confine my comments to the material discussed in Section 4.

(a) If $m=1/\rho$ is a positive integer and mY_j is a random variable which is equally likely to take the values 0, 1, . . ., $m-1$, then the autoregressive process (4) determines the conventional m-adic expansion of X_n given in equation (5) and X_n is uniformly distributed on $[0, 1[$. However, if mY_j takes the values $1-r, 2-r, \ldots, m-r$ with distinct probabilities then the m-adic expansion of X_n with these digits is still given by equation (5) but it now will have a non-uniform distribution on $[-r/m, (m-r)/m[$.

(b) Professor Bartlett stresses the fact that stochastic processes are most unlikely to be invertible. But in the special case discussed in the paper he obtains the inverse relationship.

$$x_n = mx_{n+1} - [mx_{n+1}] \tag{6a}$$

or

$$x_n \equiv mx_{n+1} \qquad (\text{mod } 1) \tag{6b}$$

which takes the form of a multiplicative congruential pseudorandom number generator. These generators are commonly used as the basis of Monte Carlo simulation studies. In the present case this procedure sequentially removes the successive leading digits of the conventional m-adic representation of x_N.

(c) To avoid the inevitable loss of accuracy which follows from an inaccurate representation of the initial value x_N, it is usually taken to be rational, and the successive values of $x_n = q_n/p$ stored as pairs of integers (q_n, p).

(d) Although it is necessary to remind oneself from time to time that pseudorandom number sequences are not truly random, and although it is somewhat worrying that such sequences may possess properties not exhibited by the random variables that they are supposed to represent, there does not seem to be any suitable alternative basis for Monte Carlo studies. In the last analysis even the sporadic discharges from thermionic valves are strictly deterministic.

Professor C. W. J. Granger (University of California, San Diego, La Jolla): A major objective of a statistical analysis is to find an approximation of the data generation process. This process may involve stochastic inputs or 'white chaotic' inputs, i.e. inputs generated by a deterministic formula that has white noise properties, such as zero autocorrelations and a flat spectrum. However, many of these white chaotic processes are quite easily seen not to be independent and identically distributed (IID). An IID series x_t has the characterizing property that $g(x_t)$ is also IID for any well-behaved function $g(\cdot)$. However, data generated from the logistic map, given by equation (3) in Professor Bartlett's paper with $k=4$, have the property that x_t has a flat spectrum but x_t^2 has a first autocorrelation of -0.22, based on a sample of 6000, which is significantly different from zero. Of course, many of the random number generators used on computers are high dimensional chaotic processes with many of the properties of IID sequences.

Brock, Deckert and Scheinkman (see Brock and Sayers (1988)) have devised a test where H_0: x_t is IID and designed to have power against an alternative of white chaos. Lee *et al.* (1989) found from a Monte Carlo study that this test had good power compared with a range of other tests against several non-linear stochastic processes, although it did not dominate a neural network test.

The possibility of chaos inputs widens the possible data-generating processes that have to be considered, although I doubt whether this is likely in the decision sciences, such as economics.

The process discussed in Section 4 of the paper is intriguing, but I would prefer a different conclusion. Reversing equation (6) indicates that there is a particular function of x_{n+1} which is perfectly forecastable from x_n, even though x_{n+1} is not perfectly forecastable. This interpretation is based on what I think is a generally accepted property of data-generating processes that the future cannot cause the present.

Dominique Guégan (Université Paris XIII): As a statistician, I agree with Professor Bartlett on the importance of chaos in model building. In modelling we need—even if the techniques are not the same—to consider chaotic models, on the same level as other stochastic models, but we have to be careful not to discern chaos everywhere.

If we consider a general definition for the chaotic model, we can say, for example, that a process $X(t)$ defined on a boundary space as (0, 1) is chaotic if, when the process has run for a long time from a nearby starting value, the process stays only in a subset of (0, 1), called the attractor set, which can be very complicated, and which has the fractional Hausdorff dimension. In fact, an important preoccupation of researchers is to catch, with chaotic models, certain non-linearities that the non-linear stochastic models do not. So, if we follow the previous definition of chaotic models, we know that the solution is linked with the dimension of the attractor set. But we know too that this approach is very difficult. So, another way to approach the problem is to consider a chaotic model such as equation (3) and to disturb it with the help of a very small random variable to approach the non-linearities that we want to identify (we need to check this random variable very carefully because we want in this case to stay very close to the original chaotic models). This approach is very nearly one which tries to find a distribution which seems stochastic but which is in fact deterministic. This problem can first be considered with the help of simulations. We can see, for example, that if we consider the joint distribution of different models as chaotic models, bilinear models or ARCH models (Engle, 1982), in certain cases they are not very dissimilar.

Another point of view, which seems to be very promising, is to establish powerful tests which allow us to decide between deterministic and stochastic models. Some work has already begun in this area, such as that of Brock *et al.* (1986).

Dr W. K. Li (University of Hong Kong): This is a stimulating paper. An important related concept is that of a self-similar process. The simplest statistical (probabilistic) model that exhibits self-similarity is that of the fractional Brownian motion introduced by Mandelbrot and Van Ness (1968). This motivated the fractionally integrated autoregressive moving average models in Granger and Joyeux (1980) and Hosking (1981). These are stationary models. Some recent empirical work by Diebold and Rudebusch (1989) on economic data provides evidence for non-stationary fractionally integrated models. Their work may furnish a possibility for the kind of research suggested by Professor Bartlett in his conclusion.

I find the comment on possible robust prediction of chaotic systems difficult to grasp as it represents some internal contradiction within the statement itself. Would Professor Bartlett elaborate on this point?

An important and challenging problem in theoretical physics is to provide a unified theory for the four fundamental forces of nature, namely the weak force, the strong force, the gravitational force and the electromagnetic force. Analogously we also have four important models in time series analysis, namely non-linear models, long memory models, non-Gaussian models and non-stationary models. A unification of these models is now possibly within reach using the idea of dynamical systems.

Professor Jian Liu (University of British Columbia, Vancouver): Professor Bartlett has pointed out the relevance of stochastics and chaos which I agree with very much. My comments are as follows.

On predictability, it is always desirable to formulate 'robust predictions' for an appropriate set of neighbouring initial conditions when possible. There are certainly cases when such an approach is not appropriate. For example, in analysing stock-exchange-type data, many considered that they are unpredictable. However, we may wish to predict some components which carry a significant amount of the total variation. This then relates closely to the concept of mean reversion in the theory of finance. Similarly, we may consider singling out some chaotic (or random) phenomena which can be 'robustly predicted'.

Although it is certainly feasible in theory as well as in certain practical situations to adopt the approach of multivariatizing a time series, owing to insufficient supply of data we may experience the same difficulty as in splining. Hence, the approximation using finitely parameterized models plays a more important role. This gives rise to an important problem of stability (or stationarity) which is one of the main characters governing predictability. As shown by many researchers (e.g. Tong (1990) and Liu and Brockwell (1988)), it is sometimes difficult to interpret the associated Lyapunov exponent in a practically usable way. Nevertheless, it provides a guideline for a search of stability conditions for a dynamical stochastic system.

Professor Murray Rosenblatt (University of California, San Diego, La Jolla): It is a pleasure to contribute a few words on Professor Bartlett's paper. He has shown great insight in his researches throughout the years on many topics. The term chaos that refers to erratic behaviour of trajectories traced out by some deterministic transformations has become popular in this sense recently because of the stimulating paper of Lorenz (1963). He showed that the trajectories of the simple non-linear system

$$\dot{x} = -10 + 10y,$$
$$\dot{y} = rx - y - xz,$$
$$\dot{z} = -\frac{8}{3}z + xy$$

can, for a range of values of r, have an erratic behaviour. This arose as an approximation in the context of convection in fluid mechanics and further applications of chaos have been proposed for many systems (see Ruelle (1989a)).

An example that I once mentioned to Bartlett belongs to a sort of ancestral folklore but exhibits many of the features that have been mentioned in later discussions of chaos. If we look at the model $X_n = \frac{1}{2}X_{n-1} + Y_n$ with the Y_n independent random variables taking on the values zero and unity with probability $\frac{1}{2}$, there is a unique stationary process $\{X_n\}$ satisfying the system of equations and it has a uniform instantaneous distribution for X_n. The forward best predictor of X_n in mean square given the past is $X_n^* = \frac{1}{2}X_{n-1}$. The backward best predictor of X_n in mean square though is $X_n^{**} = 2X_{n+1}$ modulo 1 (non-linear deterministic) and the prediction error is zero as can be noted by remarking on the binary expansions $X_n = \cdot Y_n Y_{n-1} \ldots$. The transformation $z_1 = f(z_0) = 2z_0$ modulo 1 can be regarded as a continuous transformation on the circle $0 \leqslant z_0 < 1$ where 1 is identified with 0. As already noted the uniform distribution is invariant relative to the transformation f, i.e. if z_0 is uniformly distributed on $[0, 1)$ then so is z_1. We can ask for other invariant distributions. Many can be obtained in the following way. Consider a fixed initial point z_0 and generate the distribution $F_n(\cdot, z_0)$ with mass $1/n$ at the points $f^{(j)}(z_0)$, $j = 1, \ldots, n$, where $f^{(j+1)}(z) = f\{f^{(j)}(z)\}$. Any distribution $F(\cdot)$ that is a weak limit point of the sequence $F_n(\cdot, z_0)$ is invariant relative to the transformation f (see Rosenblatt (1971)). Also if we take any stationary dependent 0, 1 sequence $\{Y_n\}$ and consider the distribution of $\cdot Y_n Y_{n-1} \ldots$ it will provide an invariant distribution relative to f.

W. M. Schaffer (University of Arizona, Tucson): Here I comment on three aspects of Professor Bartlett's presentation.

Adding noise to the SEIR model

In Schaffer (1985), the state variables were perturbed with multiplicative Gaussian noise, i.e. $X(t)$ was replaced by $X(t)\{1 + s\,Z(t)\}$. Here s is the standard deviation and $Z(t)$ samples the normal distribution. Since then, I have tried parametric excitation, whereby one or more of the parameters, for example the birth rate m, is allowed to fluctuate. In all cases, I obtained the same result: adding noise yields sustained fluctuations but does not reproduce the salient feature of the data which is the return map, which indicates that there is some predictability to the sequence of orbital excursions—in this case, the year-to-year fluctuations in case rates.

Levels of seasonality

Estimating the level of seasonality from biological grounds is difficult. In addition, the SEIR model undergoes an abrupt transition to large amplitude chaos over small parameter ranges. Moreover, as Bartlett observes, equations (1) omit all manner of relevant detail. Therefore, we need to estimate the effective seasonality, i.e. the level which, when plugged into the simplified model, gives the same behaviour that we would observe for the real value in the real equations. Using Monte Carlo methods, Lars Olsen and I have tried the following. For different parameter values and multiple trials, we estimate the correlation dimension. For large communities, the data can only be described by positing a level of seasonality well into the chaotic region. Lower levels of seasonality predict a dimension substantially in excess of what we observe in cities with populations exceeding 250 000. The calculations agree with Bartlett's own work on small populations. His 'critical population size' shows up as a sharp break in the predicted dimension.

The curious example

One of the interesting properties of chaotic motion is that we can sometimes show that the dynamics are equivalent to a sequence of symbols, e.g. 0 and 1. Thus, for the logistic map, we can subdivide the interval and describe the sequence of iterates in terms of whether a point lands in the left subinterval or the right. Interestingly, the numbers that obtain are indistinguishable from the output of a Bernoulli process. Despite this, using the numerical output, as opposed to the 0s and 1s, we can verify by plotting

$X(i+1)$ against $X(i)$ that what is generating the numbers is a deterministic map. This signature is not to be found in Professor Bartlett's example. On the contrary, each $X(i)$ can map to one of three values, and there is no information, either in the model or the time series, to tell which way the system will jump, i.e. we can mimic the statistics, but not the short-term determinism of a low dimensional chaotic process. Making the distinction may require more information than is available.

Dr Andrew Walden (BP Exploration Co. Ltd, London): I should like to elaborate a little on chaos and time series, which is currently a topic of great interest in geophysics and geology. To extract geometric information from a time series both time derivatives $x_1(t)=x(t)$, $x_2(t)=x'(t)$, $x_3(t)=x''(t)$, . . . and time delays $x_k(t)=x\{t+(k-1)h\}$, $k=1, \ldots, N$, have been proposed (Ruelle, 1989b). Concentrating on the latter, the dynamical trajectories in this N-dimensional space have the same topological characteristics as the trajectories of the system in its phase space. N must be chosen to be about twice the Hausdorff dimension (what Mandelbrot calls the fractal dimension) of the true attractor.

Most practical applications use discrete time implementations. Scargle (1989) points out that with an infinite amount of noise-free data the values of the lags do not matter as long as they are non-zero; however, in practice numerical techniques exist for determining h and/or N (e.g. Farmer *et al.*, 1983). Scargle managed successfully to deconvolve chaotic time series. Just as for a stationary random process where minimum entropy deconvolution uses an objective function which seeks to minimize disorder of estimated innovations in an entropy sense, so for a chaotic process (a filtered chaotic innovations sequence) Scargle used an objective function which minimizes the scattering of the estimated innovations in the embedding space. This objective function was also found able to deconvolve a mixture of an autoregressive random component and a chaotic process with innovations generated according to the logistic map (3).

In oil and gas exploration a simple two-dimensional space has been used by Goupillaud and Gustafson (1989) to produce a quasi-phase-space representation of a seismic trace to extract extra discriminatory information.

Broomhead and Jones (1989) used multivariate time series created from time delays of electronic oscillator data and carried out a principal components transformation on the N-dimensional samples; allowance could thus be made for quantization noise by using only the $d<N$ principal components corresponding to eigenvalues above a noise threshold. The effective embedding space then becomes the subspace spanned by the first d eigenvectors.

The **author** replied later, in writing, as follows.

My indebtedness to all the discussants must include my thanks for the many further references that they have listed. If there were any doubts about the relevance of my paper for Fellows, surely these various comments and references have dispelled them; indeed, I was gratified at the meeting to realize the amount of interest already being shown by statisticians. My own further remarks will try to co-ordinate some of this discussion, though clearly, with such an active and wide area of research, it is far too early to attempt any final assessment.

I was glad to receive comments from both Professor Schaffer and Professor Dietz (the latter with his colleague D. Schenzle) on the field of application discussed in my paper, epidemiology. These, considered together, perhaps emphasize that the rather complicated picture which I suggested is not so unreasonable. For example, Schaffer mentioned the rather abrupt transition to chaos with changes in the value of the seasonal factor, for which only an 'effective value' can be estimated, whereas Dietz and Schenzle stressed that any claim to 'excellent agreement' of the model is very dependent on the particular aspects of the data chosen for comparison (a problem arising with all model building). Other applications mentioned in the discussion include the analysis of electroencephalograms (A. Babloyantz) and geophysical time series (A. Walden).

The proposer and seconder have raised various methodological points which I can hardly cope with adequately. For example, Professor Tong referred to his new book, *Non-linear Time Series*, which I have not yet had an opportunity to look at. However, his claim to be able to separate a 'low dimensional attractor' from a 'stochastic component' or 'high dimensional attractor' seems important. (A. Walden also referred to the 'deconvolution' of chaotic and random components, and D. Guégan to tests discriminating between stochastic and deterministic models.) I do not know the answer to Tong's query about the possibility of chaos in bilinear models with *constant* parameter values, the chaos in the SEIR model being dependent on the seasonal factor (cf. Schaffer's comments in this discussion).

Professor Smith has emphasized statistical inference aspects, including the detection of chaos in real time series. His reference to Lyapunov numbers in connection with my 'curious example' is illuminating, though perhaps hardly needed in this particular case. In this context I am indebted to Murray Rosenblatt for his contribution, to A. J. Lawrance for the further Rosenblatt references and to Tong for his alternative reference to Whittle (1963) which I regret overlooking (Whittle refers in turn to Moran!). With this type of example I also overlooked the possibility noted by Lawrance of *negative correlation*. This example was also referred to by A. R. Thatcher, whose remarks on information theory in relation to chaotic systems are relevant, even if the question of 'time's arrow' is not one that I want to take up here. Smith also commented on the *dimension* of an attractor (cf. Tong's remarks), but there may be problems with this concept. For example, Osborne and Provenzale (1989) refer to a class of stochastic processes with power law spectra which give a 'fractal dimension' similar to that arising from a strange attractor. I should interpolate that my paper did not mention the important concept of *fractals*, related to, but not to be confused with, *chaos*. A *stochastic* example is white noise in continuous time. Wolf *et al.* (1985) have noted that, strictly speaking, fractal structure cannot exist either in nature (where it is truncated at the atomic level) or in any computer representations (where finite precision truncates scaling).

Subba Rao, with his colleague J. Yuan, drew attention to the useful class of bilinear models, and also to the possible value of higher order spectral analysis in studying chaotic time series. The correlation integral was mentioned by R. C. Wolff as another possible statistical tool. Dr Renshaw referred to his new book, which is another publication I have yet to see, but his remarks on the collapse of finer structures in deterministic systems in the presence of stochasticity, and his concluding comment, echoed my own plea for further study of combined stochastic and non-linear deterministic systems. A. J. Mayne noted the possible further stochasticity arising from variability in the parameters; he referred also, however, to 'fuzziness', which I believe is a new technical concept, but not one that I can claim familiarity with.

Professor Ozaki stressed not only the sensitivity of chaotic systems to initial conditions but also the possible dependence of the chaotic phenomena on a particular set of parameter values (as remarked by Schaffer for the epidemiological example, or noted in Section 3 of my paper for the discrete time logistic model). Ozaki seemed rather to dismiss 'deterministic prediction'; but at the present time predictability properties are very much to the fore in the literature, and may well play an increasingly important role in the detection and classification of chaotic systems. They are one of the topics (query (d)) among Sir David Cox's list of queries. A very recent publication is Sugihara and May (1990), though it is not clear that those researchers are able to cope with more than minor 'noise'. Cox's last query on the possibility of chaos in point processes is intriguing, but I do not see that his example of the superposition of a 'small number of strictly periodic signals' is very relevant; chaotic behaviour is manifested in the *continuous* component of the spectrum, and this is why orthodox (second-order) spectral analysis may not discriminate between chaos and stochasticity (see my 'curious example').

Predictability problems were also discussed by F. Drepper, who referred to the usefulness of an 'unpredictability profile', J. Liu, W. K. Li and by D. A. Elston and C. A. Glasbey. By 'robust prediction', to reply to Li's query, I had in mind prediction not dependent on an *impractically* precise knowledge of the initial conditions (again, my 'curious example' illustrated this), but Li should refer further to Farmer and Sidorowich (1988) and Sugihara and May (1990). In their comments Elston and Glasbey mentioned several other topics, including the problem of quantum physics. I do not want to digress on this long-standing problem unduly, but, whatever the possible relevance of chaos theory, reiterate my point about the uncertainty principle (which even precludes orthodox quantum theorists from permitting specific *phase space* trajectories). In this same context, I do not know on what basis Dr Farebrother claims that thermionic valve discharges are 'strictly deterministic'.

I have to apologize to Professor Durbin for forgetting his 1987 reference to chaos in Section 4.2 of his Presidential Address, especially as his outlook and my own on the nature of probability seem so close. Professor Granger doubted the occurrence of chaos in economics, but I should perhaps remind him that with epidemic series it was *seasonality* that played an important role.

At the meeting I suggested that even the definition of *chaos* did not yet seem resolved. The sensitivity of the non-linear system to initial conditions is one essential prerequisite (cf. Ruelle (1990)). However, a tendency to refer to the type of phenomenon referred to in my Section 5, when *transients* stop at one among more than one equilibrium point, as chaotic (rather than random) seems confusing and unnecessary; I would prefer to restrict the concept of chaos to systems which exhibit *sustained* changes (in Ruelle's sense; see also Farmer and Sidorowich (1988), p. 6), and thus to link it with 'strange attractors'.

Finally, the discussion has complemented my brief paper to produce a rather more satisfactory compilation of some of the many aspects of chaos; and, to repeat, I have been encouraged to see the active participation of statisticians in studying its interaction with more orthodox statistical theory. However, bearing in mind that probability and chance have been studied for over 300 years, I would surmise that we shall wait another 50 years or more before we can expect any reasonably complete resolution of all the new problems that have arisen. Perhaps by then we shall also see realized Professor Dawid's vision of an all-embracing theory covering probability, randomness and chaos.

REFERENCES IN THE DISCUSSION

Barnsley, M. (1988) *Fractals Everywhere*. Boston: Academic Press.

Brock, W. A. (1986) Distinguishing random and deterministic systems: a bridged version. *J. Econ. Theory*, **40**, 168–195.

Brock, W. A., Dechert, W. D. and Scheinkman, J. A. (1986) A test for independence based on the correlation dimension. *Technical Report*. University of Wisconsin, Madison.

Brock, W. A. and Sayers, C. L. (1988) Is the business cycle characterized by deterministic chaos? *J. Mon. Econ.*, **22**, 71–90.

Brockett, R. W. (1977) Convergence of Volterra Series on infinite intervals and bilinear approximations. In *Nonlinear Systems and Applications* (ed. V. Lakshmikanthan), pp. 39–46. New York: Academic Press.

Broomhead, D. S. and Jones, R. (1989) Time-series analysis. *Proc. R. Soc. Lond.* A, **423**, 103–121.

Broomhead, D. S. and King, G. P. (1986) Extracting qualitative dynamics from experimental data. *Physica* D, **20**, 217–236.

Casdagli, M. (1989) Nonlinear prediction of chaotic time series. *Physica* D, **35**, 335–356.

Chernick, M. R., Daley, D. J. and Littlejohn, R. P. (1988) A time-reversibility relationship between two Markov chains with exponential stationary distributions. *J. Appl. Probab.*, **25**, 418–422.

Cox, D. R. and Smith, W. L. (1953) The superposition of several strictly periodic sequences of events. *Biometrika*, **40**, 1–11.

Cutler, D. and Dawson, D. A. (1989) Estimation of dimension for spatially distributed data and related limit theorems. *J. Multiv. Anal.*, **28**, 115–148.

Dawid, A. P. (1985) Calibration-based empirical probability. *Ann. Statist.*, **13**, 1251–1273.

DeGroot, M. H. (1986) A conversation with Persi Diaconis. *Statist. Sci.*, **1**, 325.

Deissler, R. J. and Farmer, J. D. (1990) Deterministic noise amplifiers. Submitted to *Physica* D.

Diebold, F. X. and Rudebusch, G. D. (1989) Long memory and persistence in aggregate output. *J. Mon. Econ.*, **24**, 189–209.

Drepper, F. (1988) Unstable determinism in the information-production-profile of an epidemiological time series. In *Ecodynamics, Research Reports in Physics* (eds W. Wolff *et al.*) New York: Springer.

Drepper *et al.* (1990) The unpredictability profile of measles incidence patterns. *Preprint*.

Durbin, J. (1987) Statistics and statistical science. *J. R. Statist. Soc.* A, **150**, 177–191.

Eckmann, J.-P., Oliffson Kamphorst, S., Ruelle, D. and Ciliberto, S. (1986) Lyapunov exponents from time series. *Phys. Rev.* A, **34**, 4971–4979.

Engle, R. F. (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, **50**, 987–1007.

Falconer, K. (1990) *Fractal Geometry*. Chichester: Wiley.

Farmer, J. D., Ott, E. and Yorke, J. A. (1983) The dimension of chaotic attractors. *Physica* D, **7**, 153–180.

Farmer, J. D. and Sidorowich, J. J. (1987) Predicting chaotic time series. *Phys. Rev. Lett.*, **59**, 845–848.

——— (1988) Exploiting chaos to predict the future and reduce noise. In *Evolution, Learning and Cognition* (ed. Y. C. Lee), p. 277. Singapore: World Scientific.

Fine, P. E. M. and Clarkson, J. A. (1982) Measles in England and Wales: I, An analysis of factors underlying seasonal patterns. *Int. J. Epidem.*, **11**, 5–17.

Gaver, D. P. and Lewis, P. A. W. (1980) First-order autoregressive gamma sequences and point processes. *Adv. Appl. Probab.*, **12**, 727–745.

Goupillaud, P. L. and Gustafson, E. P. (1989) Use of chaotic dynamics and fractal geometry in geophysical interpretation. In *Expanded Abstr. 59th Meet. Society of Exploration Geophysicists, Dallas*, pp. 776–779. Society of Exploration Geophysicists.

Granger, C. W. J. and Joyeux, R. (1980) An introduction to long-memory time series models and fractional differencing. *J. Time Ser. Anal.*, **1**, 15–29.

Grassberger, P. and Procaccia, I. (1983) Characterization of strange attractors. *Phys. Rev. Lett.*, **50**, 346–349.

Hosking, J. R. M. (1981) Fractional differencing. *Biometrika*, **68**, 165–176.

Kolmogorov, A. N. and Uspenskii, V. A. (1987) Algorithms and randomness. *Theory Probab. Appl.*, **32**, 389–412.

Lee, T.-H., White, H. and Granger, C. W. J. (1989) Testing for neglected non-linearity in time series models: a comparison of neural network methods and alternative tests. *Discussion Paper 89-42*. Economics Department, University of California, San Diego, La Jolla.

Liu, J. and Brockwell, P. J. (1988) On the general bilinear model. *J. Appl. Probab.*, **25**, 553–564.

London, W. P. and Yorke, J. A. (1973) Recurrent outbreaks of measles, chickenpox and mumps: I, Seasonal variation in contact rates. *Am. J. Epidem.*, **98**, 453–468.
Lorenz, E. N. (1963) Deterministic nonperiodic flow. *J. Atmosp. Sci.*, **20**, 130–141.
Mandelbrot, B. B. and Van Ness, J. W. (1965) Fractional Brownian motions, fractional noises and applications. *SIAM Rev.*, **10**, 422–437.
Osborne, A. R. and Provenzale, A. (1989) Finite correlation dimension for stochastic systems with power-law spectra. *Physica* D, **35**, 357–381.
Prigogine, I. and Stengers, I. (1986) *Order out of Chaos*. London: Fontana.
Ramsey, J. B. and Yuan, H.-J. (1990) The statistical properties of dimension calculations using small data sets. *Nonlinearity*, **3**, 155–176.
Renshaw, E. (1990) *Modelling Biological Populations in Space and Time*. Cambridge: Cambridge University Press.
Rissanen, J. (1989) *Stochastic Complexity in Statistical Enquiry*. Singapore: World Scientific.
Rosenblatt, M. (1971) *Markov Processes: Structure and Asymptotic Behavior*, p. 101. New York: Springer.
——— (1979) Linearity and non-linearity in time series. In *Proc. 42nd Int. Statist. Inst., Manila*, book 1, pp. 423–434. International Statistical Institute.
——— (1980) Linear processes and bispectra. *J. Appl. Probab.*, **17**, 265–270.
——— (1985) Prediction and moments (problem 10, comment 2.4). In *Stationary Sequences and Random Fields*. Boston: Birkhauser.
——— (1986) Prediction for some non-Gaussian autoregressive schemes. *Adv. Appl. Math.*, **7**, 182–198.
Ruelle, D. (1989a) *Elements of Differentiable Dynamics and Bifurcation Theory*. New York: Academic Press.
——— (1989b) *Chaotic Evolution and Strange Attractors*, p. 28. Cambridge: Cambridge University Press.
——— (1990) Deterministic chaos: the science and the fiction (the Claude Bernard Lecture 1989). *Proc. R. Soc. Lond.* A, **427**, 241–248.
Sakai, H. and Tokumaru, H. (1980) Autocorrelations of a certain chaos. *IEEE Trans. Acoust. Spch Sig. Process.*, **28**, 588–590.
Sano, M. and Saweda, Y. (1985) Measurement of the Lyapunov spectrum from a chaotic time series. *Phys. Rev. Lett.*, **55**, 1082–1085.
Scargle, J. D. (1989) An introduction to chaotic and random time series analysis. *Int. J. Imgng Syst. Technol.*, **1**, 243–253.
Schaffer, W. M. (1985) Can nonlinear dynamics elucidate mechanisms in ecology and epidemiology? *IMA J. Math. Appl. Med. Biol.*, **2**, 221–252.
Schenzle, D. (1984) An age-structured model of pre- and post-vaccination measles transmission. *IMA J. Math. Appl. Med. Biol.*, **1**, 169–191.
Sesay, S. A. O. and Subba Rao, T. (1988) Yule–Walker type difference equations for higher order moments and cumulants for bilinear time series models. *J. Time Ser. Anal.*, **9**, 385–401.
Smith, L. A. (1988) Intrinsic limits on dimension calculations. *Phys. Lett.* A, **133**, 283–288.
Subba Rao, T. (1981) On the theory of bilinear time series models. *J. R. Statist. Soc.* B, **43**, 244–255.
Subba Rao, T. and Gabr, M. M. (1984) An introduction to bispectral analysis and bilinear time series models. *Lect. Notes Statist.*, **24**.
Sugihara, G. and May, R. M. (1990) Non-linear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature*, **344**, 734–741.
Taylor, C. C. and Taylor, S. J. (1991) Estimating the dimension of a fractal. *J. R. Statist. Soc.* B, to be published.
Tong, H. (1990) *Non-linear Time Series: a Dynamical System Approach*. Oxford: Oxford University Press.
Vovk, V. G. (1987) *Theory Probab. Appl.*, **32**, 413–425.
Whittle, P. (1963) *Prediction and Regulations*, sect. 2.6. Oxford: Blackwell.
Wolf, A., Swift, J. B., Swinney, H. L. and Vastano, J. (1985) Determining Lyapunov exponents from a time series. *Physica* D, **16**, 285–317.
Wolff, R. C. (1990) A note on the behaviour of the correlation integral in the presence of a time series. *Biometrika*, to be published.
Yule, P. (1927) On the method of investigating periodicities in distributed series with special reference to Wolfer's sunspot numbers. *Phil. Trans. R. Soc. Lond.* A, **226**, 267–298.